forthcoming in Behavioral and Brain Sciences
[Pre-edit draft: please cite final published version]

BBS Commentary on Susan Carey, *The Origin of Concepts*

Acquiring a new concept is not explicable-by-content

Nicholas Shea

Abstract
Carey's book describes many cases where children develop new concepts with expressive power that could not be constructed out of their input. How does she side-step Fodor's paradox of radical concept nativism? I suggest it is by rejecting the tacit assumption that psychology can only explain concept acquisition when it occurs by rational inference or other transitions that are explicable-by-content.

Main text
Representational explanation is central to psychology. Mental processes are characterised in terms of causal transitions between token states, where we make sense of the transitions in terms of the content of the states. Can we explain concept acquisition in the same way? Only insofar as the acquired concept is constructed out of pre-existing concepts, according to Fodor. Since most lexical concepts do not seem to be so-constructed, Fodor concludes that they are innate – their acquisition is outside the ambit of psychological explanation.

Carey's book is a comprehensive refutation of radical concept nativism, offering many psychological explanations of concept acquisition, and the data to back them up. How, then, does Carey side-step Fodor's argument? I want to suggest that she has to reject the assumption that all psychological explanations are explanations-in-virtue-of-content. Consider two of the stages in Carey's account of the acquisition of number concepts.

First, the transition from parallel individuation to enriched parallel individuation (being a one-knower, two-knower, etc.). Numerosity is not represented explicitly anywhere in the parallel individuation system. It is implicit in the various operations that are performed on object files: adding, subtracting and comparing by 1-1 correspondence. The child then comes to associate words with object files of a certain size, e.g. "one" with having one object file of any kind open: {i}. Is this step explicable-by-content?

Before becoming a one-knower, the child was not representing one-ness explicitly at all: "one" was just a sound, and numerosity was merely implicit in the object file system. The child did not have resources out of which a hypothesis about one-ness could be constructed. But there were two important correlations that they could make use of: (1) between having one object file {i} open and singleton sets; and (2) between the word "one" and singleton sets (the mechanism for which involves the child's linguistic community). Although neither is a representation of one-ness, these are two pieces of *information*, of the purely correlational type (e.g. Shannon information). Since the two mental items correlate with the same external-world property, they tend to occur together, so become associated. The association between

"one" and {i} constitutes a new symbol. It explicitly represents the numerosity *one* (i.e. that is its wide content).

I would argue that that transition has not been explained-by-content. Instead is a transition to an entirely new representation, explained in terms of the correlational information carried by its precursors. One of those precursors ("one") was not representational at all and the other ({i}) was not made use of for its content, but just for its informational connections. The transition to the new symbol ("one"/{i}) was no kind of inference or other rational transition from those pre-existing resources. Nevertheless, Carey has offered a recognisably psychological explanation. In my view this shows that there can be psychological explanations in terms of correlational information that are not explanations-by-content.

For a second example, consider the transition from enriched parallel individuation to being a cardinal principle knower. This step involves Quinean bootstrapping, the process whereby a set of uninterpreted symbols interrelated by a network of inferential dispositions are connected up to the world so as to acquire a meaning. The child's key resource here is the uninterpreted list of counting words: "one", "two", … . At the stage of enriched parallel individuation the early words in this sequence have already been put into correlations with object files ("one" / {i}), hence numerosities (one-ness). To give the symbols from "five" onwards their content, according to Carey, the child generalises across three transitions in which moving to the next count word corresponds to adding one to the object file:

"one"/{i} → "two"/{i j}

"two"/{i j} → "three"/{i j k}

"three"/{i j k} → "four"/{i j k l}

The child makes a leap which generalises across the instances of adding one – by associating them all as instances of counting on. When he does, "five" is put into a content-constituting informational connection with sets of five things – kept track of as the successor to four, which the child can track directly with the enriched object file system using its symbol "four"/{i j k l} (similarly for "six", …).

Is this transition explicable-by-content? Carey rightly rejects the idea that it is properly described as hypothesis testing. The child is doing something different: building an uninterpreted model with the counting words, and then giving that model an interpretation. He does not test a hypothesis (a statement formulable with his existing representational resources). Rather, he comes to associate two previously-correlated operations (counting on and adding one). By doing so, he acquires the concept of successor (generalising over instances of adding one) and concepts of all the numbers for which he has count words. Again, Carey has described a psychological process that depends upon correlational information but is not explicable as a rational inference or other transition-in-virtue-of-content.

In some places Carey suggests that Quinean bootstrapping is explicable-by-content, because even before the transition the uninterpreted placeholders have narrow contents in virtue of their inferential roles (p. 522). Carey has some good arguments for the existence of

narrow content, but not in the case of dispositions to make transitions between uninterpreted symbols. The transitions are described in terms of connections between symbol types, where the symbols can only be individuated non-semantically. Once the symbols are put in the right relations to acquire wide contents, then we can use inferential dispositions between concepts to characterise a second level of narrow content. However, on pain of regress or holism, the inferences that make up the narrow content of a concept should be individuated in terms of the wide contents of the concepts which figure in those inferences. So narrow contents cannot save explanation-by-content.

With hindsight, it is obvious that if we assume that psychological explanation is restricted to explanation-by-content, then psychology is going to have a problem explaining the acquisition of genuinely new representational resources, because the required contents would have to be available before the transition. Carey's book gives us compelling reasons for relinquishing that assumption.

N.S.
21 June 2010